Regional characteristics and the decision to innovate in a developing country: A multilevel analysis of Ecuadorian firms

Supplementary Appendix

Bruna, F. and Fernández-Sastre, J. (2021): Regional characteristics and the decision to innovate in a developing country: A multilevel analysis of Ecuadorian firms. *Papers in Regional Science* 100(6), 1337-1354. https://doi.org/10.1111/pirs.12632

1. Observations removed

ENAI only has coverage for firms with more than 10 employees. However, in the database there are firms reporting less than 10 employees in the three years of the survey. This might be due to a lag between the sample selection and the gathering of the information. We have kept those firms in our sample but removed the following 153 observations:

- 17 firms with zero sales in the three years of the survey
- 26 firms with employment zero in the three years of the survey
- 52 firms with average R&D investment in the period 2012-2014 less than 100 USD
- 5 firms with average employment in the period 2012-2014 less than 1
- 53 firms that did not report information on the level of qualification of employees

2. The multilevel logit model

We follow James et al. (2017, chap. 4) for general description of the logistic model and Snijders and Bosker (2012, 290) for multilevel notation.[1] This way of presenting the multilevel logistic model is similar to what Srholec (2010; Srholec 2011) and Barasa et al. (2017) do[2] For firm i (level 1) in province j (level 2, $j = 1, \dots, 24$), the dichotomous outcome variable Y_{ij} is coded 1 if a firm develops innovation activities and 0 otherwise. Y_{ij} is represented as the sum of an expected value P_{ij} and a residual R_{ij}, as in the following:

$$Y_{ij} = P_{ij} + R_{ij} \tag{1}$$

[1] The Centre for Multilevel Modelling of the University of Bristol provides a comprehensive list of resources. Chapter 7 of the Centre's free course explains the logit multilevel model. See also Sommet & Morselli (2017).
[2] An alternative representation is the so-called threshold model (Snijders and Bosker 2012, 303), in which the dichotomous outcome is conceived as the result of a latent variable. This method introduces the innovation multilevel model by Bellmann et al. (2018) and López-Bazo & Montellón (2018).

where P_{ij} is the probability of innovation of firm i in region j, such that its values range from 0 to 1. The zero-mean residual R_{ij} can take only the values $-P_{ij}$ and $1 - P_{ij}$ if Y_{ij} is to be 0 or 1. We want to predict P_{ij} from a set X_{ij} of firm-level explanatory variables, allowing for *between-context heterogeneity*,[3] which accounts for provincial differences in the dependent variable. We thus obtain a random intercept multilevel model that now appears as a regional (provincial) intercept γ_j that takes 24 possible values. The following *logistic function* produces values between 0 and 1:

$$P_{ij} = \frac{\exp\left(\gamma_j + \gamma_{10} X_{ij}\right)}{1 + \exp\left(\gamma_j + \gamma_{10} X_{ij}\right)} \tag{2}$$

We seek estimates for the parameters γ_j and γ_{10}, such that the predicted probabilities $\hat{P}_{ij}$ of innovation for each firm correspond as closely as possible to the firm's observed innovation status. The estimates are chosen to maximize a likelihood function capturing the joint probability of 0 and 1 observed in the sample of Y_{ij}.[4] Indeed, the maximized function is the log-likelihood. To interpret the estimation of equation (2), consider that the logistic model has a *logit* that is linear in X_{ij}. We observe this by first manipulating equation (2) to produce *odds*, as in the left side of the following equation:

$$\frac{P_{ij}}{1 - P_{ij}} = \exp\left(\gamma_j + \gamma_{10} X_{ij}\right) \tag{3}$$

Since odds are the ratio of probability of innovation to probability of no innovation, $P_{ij}/\left(1 - P_{ij}\right)$, they can vary from zero to infinity. The natural logarithm transforms the set of positive real numbers into the whole real line. The following logit function is the log-odds for firm i in province j and models the propensity to innovate of firm i in region j:

$$logit\left(P_{ij}\right) = \ln\left(\frac{P_{ij}}{1 - P_{ij}}\right) = \gamma_j + \gamma_{10} X_{ij} \tag{4}$$

where $\exp\left(\gamma_j\right)$ represents the odds of innovating when $X_{ij} = 0$, and $\exp(\gamma_{10})$ is the odds ratio comparing probability of innovating for two firms with values of X_{ij} spaced one unit apart. An estimated odds ratio lower than 1 implies that an increase in this explanatory variable has a negative impact on probability of innovation.

[3] See Duncan et al. (1998) and Neira et al. (2018).

[4] This maximization has not closed-form solution, but an approximate solution can be found numerically.

To transform the firm-level equation (4) into a multilevel model, we need another equation that models the regional (provincial) intercept, as in the following (macro) level-2 random effects model:

$$\gamma_j = \gamma_0 + \gamma_{01} Z_j + U_j \qquad U_j \sim \mathcal{N}(0, \sigma_j^2) \tag{5}$$

where γ_0 is a global intercept for all firms, Z_j is a set of provincial variables whose values are the same for all firms in the same province, and γ_{01} is a set of parameters capturing the effect of Z_j on γ_j. U_j are the random deviations of each sector from the provincial intercepts $\gamma_0 + \gamma_{01} Z_j$. We estimate these values as provincial residuals after estimating the parameters for coefficients and variances.

Taking equations (4) and (5), we obtain the following joint random effects multilevel model:

$$logit(P_{ij}) = \ln\left(\frac{P_{ij}}{1 - P_{ij}}\right) = \gamma_0 + \gamma_{10} X_{ij} + \gamma_{01} Z_j + U_j \qquad U_j \sim \mathcal{N}(0, \sigma_j^2) \tag{6}$$

As the level-1 (firm) residuals of the binary variable Y_{ij} of innovation are given by R_{ij} in equation (1), the model in equation (6) does not include level-1 residuals. The variance of the level-1 residual error in model (4) is thus assumed to be of standard logistic distribution: $\sigma^2 = \pi^2/3 \approx 3.29$. The variance of U_j, the unexplained provincial part of the log-odds, is σ_j^2, making the variance in total error of the model $\sigma_j^2 + 3.29$. The percentage of unexplained variance due to differences among provinces is given by the following variance partition coefficient, which in this model is the same as the intra-class correlation coefficient (ICC):[5]

$$ICC = \frac{\sigma_j^2}{\sigma_j^2 + 3.29} \tag{7}$$

The higher the U_j variance, σ_j^2, the larger the unexplained variation of the average log-odds between provinces. Assuming that individual-level variance σ^2 is fixed at 3.29, the additional level-1 explanatory variable (X_{ij}) will either increase or decrease the ICC, depending on the correlation between the explanatory variables and their distribution across

[5] We do not estimate random slope models in this paper. The ICC for these models is not the same as the variance partition coefficient.

regions.[6] Adding significant level-two explanatory variables (Z_j) should reduce the ICC, however, and this is what we want to analyze.

Our first random intercepts model will be the *empty* or *null* model (Model 0), without explanatory variables: $logit(P_{ij}) = \gamma_0 + U_j$. The exponent of the regional intercepts is thus an estimate of the average provincial odds.[7] These intercepts are an indicator of the role of regional factors in shaping firm innovation decisions.

3. Analysis by sector

The Ecuadorian Innovation Survey (ENAI) contains detail for each firm about the International Standard Industrial Classification (ISIC) classification of industries at two aggregation levels: for 14 and for 63 industries. To test possible common technological factors between industries, we have aggregated the 63 industries in 9 groups, shown in Table A1. Table A2 shows the composition of the sample by each of these sectors, attending to the sectoral classification of technological intensity. Table A3 shows the estimates of the sectoral fixed effects in our preferred models of Table 3, relative to the reference category, high-tech manufactures. Those estimates have no specific interpretations, given that they reflect the probability of investing in innovation activities in each sector, compared to the reference category, once all the internal and external factors are controlled for.

Table A1 –Sector classification

	ISIC 2
Low-tech manufactures	
Manufacture of food products, beverages and tobacco	C10-12
Textile industry	C13
Wearing apparel	C14

[6] The patterns are different for dichotomous and continuous dependent variables. With continuous dependent variables, controlling for different types of firms in each region (compositional effects) may reveal unexplained regional differences, increasing the ICC. However, adding level-1 explanatory variables usually reduces the ICC. See, for instance, https://www.researchgate.net/post/Why-does-the-intraclass-correlation-coefficient-ICC-increases-as-fixed-components-are-added-in-a-multilevel-logistic-regression-model.

[7] The random effects are precision-weighted residuals called 'posterior residuals', 'empirical Bayes estimates', or 'shrunken residuals'. They are estimated by considering the number of observations in each region and the dispersion of the data within and between regions.

Leather and related products	C15
Wood and products of wood	C16
Paper and paper products	C17
Printing and reproduction of recoded media	C18
Manufacture of furniture	C31
Other manufacturing	C32
Medium-low-tech manufactures	
Manufacture of rubber and plastic products	C19
Manufacture of other non-metallic mineral products	C23
Manufacture of basic metals	C24
Manufacture of basic metals	C25
Repair and installation of machinery and equipment	C33
Repair of computers and household goods	S95
Medium-high-tech manufactures	
Manufacture of chemical and chemical products	C20
Manufacture of machinery and equipment	C28
Manufacture of motor vehicles, trailers and semitrailers	C29
Manufacture of other transport equipment	C30
High-tech manufactures	
Manufacture of pharmaceutical products and preparations	C21
Manufacture of computer, electronic and optical products	C26
Non-knowledge intensive business services (nkibs)	
Support service activities for mining and quarrying	B09
Trade	G45-47
Transport and warehousing	H49-53
Real estate activities	I55, L68
Food service activities	I56
Advertising	M73
Rental and leasing activities	N77
Employment activities	N78
Tourist activities	N79

Building and landscape service activities	N81
Administrative activities and auxiliary services	N82
Other services	O84, S94, S96, T97, T98, U99
Knowledge intensive business services (kibs)	
Publishing activities	J58
Programming and broadcasting activities	J60
Telecommunications	J61-62
Other information and communication services	J63
Legal and accounting activities	M69
Consulting	M70
Architecture, engineering, and technical analysis	M71
Other professional, scientific, and technical activities	M74
Veterinary activities	M75
Security and research activities	N80
Education	P85
Human health and social work activities	Q86-88
Arts, entertainment, and recreation	R90-R93
Utilities	
Electricity, steam, and air conditioning supplies	D35
Other utilities	E35-39
Extractive industries	B05-08
Construction	F41-43

Table A2 -Number of firms in the sample by sector and province

Province	Construction	Extractive	High-tech manuf.	Kibs	Nkibs	Low-tech manuf.	Medium-high-tech manuf.	Medium-low-tech manuf.	Utilities	Total	% / Total
Azuay	33	48	27	65	242	5	34	108	8	570	9.3
Bolívar	4	0	0	10	5	0	0	4	2	25	0.4
Cañar	5	1	0	13	16	0	5	6	5	51	0.8
Carchi	2	0	0	10	42	0	1	4	1	60	1.0
Chimborazo	7	1	6	37	54	1	9	26	2	143	2.3
Cotopaxi	13	0	2	31	62	0	10	26	2	146	2.4
El Oro	29	54	9	32	147	0	9	28	6	314	5.1
Esmeraldas	5	1	2	19	48	0	7	13	2	97	1.6
Galápagos	1	0	0	2	21	0	0	0	1	25	0.4
Guayas	101	17	93	141	712	10	84	194	19	1,371	22.4
Imbabura	11	4	3	23	83	0	12	49	2	187	3.1
Loja	26	3	1	35	77	1	8	19	3	173	2.8
Los Ríos	5	0	4	13	72	0	1	27	1	123	2.0
Manabí	32	3	5	44	145	1	12	67	7	316	5.2
Morona Santiago	3	0	0	9	9	0	0	0	3	24	0.4
Napo	10	1	0	1	7	0	1	5	1	26	0.4
Orellana	9	0	0	3	43	0	7	2	1	65	1.1
Pastaza	6	1	0	3	12	0	0	5	1	28	0.5
Pichincha	113	11	98	260	827	22	66	227	17	1,641	26.8
Santa Elena	7	1	1	12	51	0	1	17	5	95	1.6
Santo Domingo	20	1	7	25	93	0	5	35	2	188	3.1
Sucumbíos	11	0	1	12	31	0	4	1	2	62	1.0
Tungurahua	16	3	29	55	137	2	12	107	5	366	6.0
Zamora Chinchipe	8	6	0	4	7	0	0	0	1	26	0.4
Total	477	156	288	859	2,943	42	288	970	99	6,122	100.0
% / Total	7.8	2.5	4.7	14.0	48.1	0.7	4.7	15.8	1.6	100.0	

Table A3 – Estimated fixed sectoral effects in Model 3 of Table 3

	(5) R&D		**(6) Other innovation**	
	Log-Odds	*p*	*Log-Odds*	*p*
Low-tech manufactures	0.83	0.056	0.45	0.301
Medium-low-tech manufactures	1.03	**0.026**	0.47	0.295
Medium-high-tech manufactures	0.76	0.091	0.12	0.795
Non-knowledge intensive business services (nkibs)	1.20	**0.006**	-0.09	0.838
Knowledge intensive business services (kibs)	0.72	0.098	0.34	0.434
Utilities	0.44	0.469	0.20	0.688
Extractive industries	1.04	0.059	-0.46	0.345
Construction	1.04	**0.024**	0.03	0.939

Note: The reference category is high-tech manufactures.

Results from Table A3 indicate that in Ecuador there are no sectoral differences in the likelihood of investing in other innovation activities. Additionally, firms belonging to medium-low-tech manufactures, non-knowledge business services and construction are more prone to invest in R&D. This reveals the lack of R&D capabilities of firms belonging to knowledge-intensive sectors in developing countries.

4. Robustness analysis for *Other innovation*

Note that the dependent variable *Other innovation* includes investments in activities which are directly involved in innovation. Section V.4 of the ENAI indicates that *Other innovation* activities, either by their very nature, or by the sense in which they are being implemented, should contribute to obtain an innovation. Note that, in the case of machinery and equipment, or hardware, they should not be understood as a simple increase in the productive capacity of the company. Additionally, given that *Other innovation* includes several different things, it is worthy to check if our results are robust to a more restrictive definition of the variables which only includes the acquisition of technology involved in innovation.

Table A4 compares the estimation results for the dependent variables *Other innovation* and *Acquisitions*. *Other innovation* includes investment in: (1) acquisition of machinery and equipment, (2) acquisition of hardware, (3) acquisition of software, (3) acquisition of unincorporated technology, (4) hiring of consultancies and technical assistance, (5) engineering and industrial design activities, (6) training of personnel, or (7) market studies. *Acquisitions* only includes investment in the acquisition of technologies for the introduction of new products or processes (1-3).

Column (1) in Table A4 is the same as equation (6) in Table 3 for *Other Innovation*, while column (2) shows the results for the dependent variable *Acquisitions*. *Skilled labor* is no longer significant in the *Acquisitions* equation, which might suggest that the qualification of the work force is especially relevant for investments in consultancies and technical assistance, engineering and industrial design, training of personnel and market studies. Additionally, the regional-level variable *Loans to GVA* is no longer significant. Otherwise, the results are very similar.

Column (3) of Table A4 shows the result of a robustness analysis for *Other Innovation*, when the explanatory variable *Capital investment* is omitted from the model. As summarized in footnote 12 of the paper, the results are very similar to those of column (1) except for *Business to total R&D*, which is no longer statistically significant.

Table A4 – Robustness analysis of equations (5) and (6) in Table 3

	(1)		(2)		(3)	
	Other innovation		**Acquisitions**		**Other innovation**	
	Log-Odds	*p*	*Log-Odds*	*p*	*Log-Odds*	*p*
Fixed effects: internal factors (X_{ij})						
(Global intercept)	-1.44	**0.001**	-1.81	**<0.001**	**-0.04**	**0.910**
Log Employment	0.16	**<0.001**	0.12	**<0.001**	**0.27**	**<0.001**
Skilled labor	0.39	**0.007**	0.19	0.194	0.37	0.004
Capital investment	2.05	**<0.001**	2.27	**<0.001**		
Fixed effects: provincial factors (Z_j)						
Loans to GVA	0.17	**0.027**	0.09	0.256	0.15	**0.004**
Business to total R&D	0.10	**0.026**	0.11	**0.017**	0.03	0.397
Cooperating firms	0.56	**<0.001**	0.54	**<0.001**	0.65	**<0.001**
Region-sector R&D spillover	0.52	**<0.001**	0.46	**<0.001**	0.55	**<0.001**
Inter-region R&D spillover	0.43	**<0.001**	0.38	**<0.001**	0.44	**<0.001**
Random Effects						
σ^2 (level 1, assumed)	3.29		3.29		3.29	
σ_j^2 (level 2, estimated)	0.02		0.02		0.00	
Intraclass correlation (ICC)	0.01		0.01		0.00	
Deviance	5,999.2		5,785.6		7,053.6	

Notes: (1) 6,122 firms (i) in 24 provinces (j). (2) The two firm-level continuous variables, the log of employment and the % of skilled labor, are considered as deviations to the provincial mean: $X_{ij} - \bar{X}_j$. The Z_j provincial factors are centered and standardized. (3) The models include sectoral fixed effects, classified by technological intensity.

5. Models 2 and 3 of Table 3 estimated by non-multilevel models

Table A5 shows the results of standard one-level logit models, using standard errors clustered by region. Focusing on Model 3, for *R&D* the results are almost identical to those of Model 3, given that the regional ICC in this latter model is zero. For *Other innovation*, the results in Model 3 of Table A5 are very similar to those of Table 3, except for *Business to total R&D,* which becomes non-significant. However, the assumption of the one-level models is that there is not remaining variance explained by level-2 variables. The ICC in Table 3 for Model 3 for *Other innovation* is 1%, which is enough to prefer the results of Table 3 instead of those in Table A5.

Table A5 – Models 2 and 3 of Table 3 estimated by standard one-level logit models

	Model 2				**Model 3**			
	(1) R&D		**(2) Other innovation**		**(3) R&D**		**(4) Other innovation**	
	Log-Odds	*p*	*Log-Odds*	*p*	*Log-Odds*	*p*	*Log-Odds*	*p*
(Intercept)	-4.16	**<0.001**	-1.43	**0.001**	-4.01	**<0.001**	-1.41	**0.001**
Log Employment	0.17	**<0.001**	0.15	**<0.001**	0.19	**<0.001**	0.16	**<0.001**
Skilled labor	0.52	**0.001**	0.42	**0.030**	0.56	**<0.001**	0.39	**0.042**
Capital investment	1.67	**<0.001**	2.04	**<0.001**	1.64	**<0.001**	2.04	**<0.001**
Merger or acquisition	0.73	**<0.001**	0.21	0.408	0.70	**<0.001**		
Foreign group	-0.69	**<0.001**	-0.17	**0.001**	-0.66	**<0.001**		
Exporter	0.16	**0.019**	0.08	0.168				
Loans to GVA	0.05	0.285	0.20	**0.024**	0.10	**0.019**	0.19	**0.026**
Business to total R&D	0.05	0.172	0.07	0.093	0.14	**0.007**	0.07	0.100
Cooperating firms	-0.15	**<0.001**	0.57	**<0.001**			0.57	**<0.001**
Region-sector R&D spillover	1.33	**<0.001**	0.52	**<0.001**	1.36	**<0.001**	0.52	**<0.001**
Inter-region R&D spillover	-0.24	**0.001**	0.48	**<0.001**			0.48	**<0.001**
Deviance	3,918.1		6,002.2		3,932.6		6,004.7	

Notes: 6,122 firms (i) in 24 provinces (j). Statistical significance calculated with standard errors clustered by province. (2) The two firm-level continuous variables, the log of employment and the % of skilled labor, are considered as deviations to the provincial mean: $X_{ij} - \bar{X}_j$. The Z_j provincial factors are centered and standardized. (3) The models include sectoral fixed effects, classified by technological intensity.

6. Table 3 for odds ratios and robustness analysis for regional per capita income

Table A6 repeats the main results of Table 3 but showing odds ratios to better compare the effect size of the regional variables. Odds ratios lower than one imply a negative effect of that variable on the probability of investing in innovation.

Additionally, Table A6 shows that the results of Table 3 are robust to the inclusion of a standard controls of the local context such as the logarithm of Gross Value Added (GVA) per capita.

Table A6 – Testing de addition of Gross Value Added per capita in Models 2 and 3 of Table 3

	Model 2								**Model 3**							
	R&D		**Other innovation**		**R&D**		**Other innovation**		**R&D**		**Other innovation**		**R&D**		**Other innovation**	
	Odds Ratios	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*
(Intercept)	0.02	**<0.001**	0.23	**0.001**	0.02	**<0.001**	0.23	**0.001**	0.02	**<0.001**	0.24	**0.001**	0.02	**<0.001**	0.24	**0.001**
Log Employment	1.19	**<0.001**	1.17	**<0.001**	1.19	**<0.001**	1.17	**<0.001**	1.20	**<0.001**	1.17	**<0.001**	1.20	**<0.001**	1.17	**<0.001**
Skilled labor	1.68	**0.006**	1.51	**0.006**	1.68	**0.006**	1.51	**0.006**	1.74	**0.003**	1.48	**0.007**	1.73	**0.004**	1.48	**0.007**
Capital investment	5.29	**<0.001**	7.77	**<0.001**	5.29	**<0.001**	7.78	**<0.001**	5.14	**<0.001**	7.77	**<0.001**	5.18	**<0.001**	7.78	**<0.001**
Merger or acquisition	2.07	**0.004**	1.22	0.383	2.07	**0.004**	1.22	0.385	2.02	**0.005**			2.02	**0.005**		
Foreign group	0.50	**<0.001**	0.85	0.234	0.50	**<0.001**	0.85	0.237	0.52	**<0.001**			0.53	**<0.001**		
Exporter	1.17	0.173	1.10	0.364	1.17	0.168	1.10	0.361								
Loans to GVA	1.05	0.503	1.18	**0.027**	1.09	0.382	1.19	**0.026**	1.11	**0.030**	1.18	**0.027**	1.20	**0.006**	1.19	**0.026**
Business to total R&D	1.06	0.355	1.10	**0.027**	1.06	0.307	1.10	**0.026**	1.15	**0.008**	1.10	**0.026**	1.16	**0.005**	1.10	**0.025**
Cooperating firms	0.86	**0.004**	1.76	**<0.001**	0.87	**0.010**	1.77	**<0.001**			1.76	**<0.001**			1.76	**<0.001**
Region-sector R&D spillover	3.78	**<0.001**	1.68	**<0.001**	3.78	**<0.001**	1.68	**<0.001**	3.89	**<0.001**	1.68	**<0.001**	3.87	**<0.001**	1.68	**<0.001**
Inter-region R&D spillover	0.78	**0.026**	1.53	**<0.001**	0.80	0.064	1.54	**<0.001**			1.53	**<0.001**			1.54	**<0.001**
Log GVA per capita					0.96	0.569	0.99	0.778					0.88	0.064	0.99	0.759
Random Effects																
σ^2	3.29		3.29		3.29		3.29		3.29		3.29		3.29		3.29	
σ_j^2	0.00		0.02		0.00		0.02		0.00		0.02		0.00		0.02	
ICC			0.01				0.01		0.00		0.01				0.01	
Deviance	3,918.1		5,996.7		3,917.8		5,.996.6		3,932.6		5,999.2		3,928.7		5,999.1	

7. Analysis of Model 3 of Table 3 by size of the firm

Table A7 shows the estimation of our preferred models for subsamples of *Small* and *Large* firms. In order to divide the sample, we use the categorical variable provided by the ENAI, which takes 3 values: 1 if the firm has from 10 to 49 employees (3,128 small firms in our sample); 2 if the firm has from 50 to 450 employees (543 medium firms); and 3 if the firm has more than 450 employees (2,451 large firms). Our subsample of *Small* firms integrates the first two categories of the variable, while the subsample of *Large* firms matches the third category.

Table A7 – Estimation of equations (5) and (6) of Table 3 by size of the firm

	R&D				**Other innovation**			
	(1) Small		**(2) Large**		**(1) Small**		**(2) Large**	
	Log-Odds	*p*	*Log-Odds*	*p*	*Log-Odds*	*p*	*Log-Odds*	*p*
Fixed effects: internal factors (X_{ij})								
(Global intercept)	-4.57	**<0.001**	-3.40	**<0.001**	-1.52	**0.034**	-1.40	**0.014**
Log Employment	0.17	**0.024**	0.27	**<0.001**	0.18	**0.001**	0.22	**<0.001**
Skilled labor	0.81	**0.005**	0.51	**0.047**	0.72	**<0.001**	0.26	0.225
Capital investment	1.74	**<0.001**	1.57	**<0.001**	1.98	**<0.001**	2.19	**<0.001**
Merger or acquisition	0.81	0.100	0.64	**0.029**				
Foreign group	-0.24	0.630	-0.57	**0.001**				
Fixed effects: provincial factors (Z_j)								
Loans to GVA	0.22	**0.005**	-0.04	0.634	0.25	**0.001**	0.13	0.266
Business to total R&D	0.14	**0.045**	0.11	0.242	0.08	0.067	0.19	**0.039**
Region-sector R&D spillover	1.53	**<0.001**	1.13	**<0.001**	0.56	**<0.001**	0.48	**<0.001**
Cooperating firms					0.54	**<0.001**	0.68	**<0.001**
Inter-region R&D spillover					0.46	**<0.001**	0.54	**0.001**
Random Effects								
σ^2 (level 1, assumed)	3.29		3.29		3.29		3.29	
σ_j^2 (level 2, estimated)	0.01		0.00		0.01		0.00	
Intraclass correlation (ICC)	0.00				0.00			
Observations	3,671		2,451		3,671		2,451	
Deviance	1,923.4		1,970.2		3,560.4		2,410.4	

Note: See the notes to Table A3.

Results reveal some regional differences between large and small firms. First, regional volume of loans (*Loans to GVA*) is only significant for small firms' innovation decisions. This is not surprising, as large firms are more likely to finance their R&D activities with internal resources, while small companies have a greater need to resort to external financing (O'Sullivan, 2005). Second, *Business to total R&D* is positively associated with small firm's R&D investment decision and with large firms' other innovation activities investment decision. This suggests that knowledge developed by research institutions has less relevance in the R&D activities of small firms.

8. Geography of the random effects estimated for Model 3 in Table 3

Tables A8 and A9 show that the level-2 residuals, which are the estimated random effects, have an approximately Gaussian distribution. In this section, we study if these residuals have a random spatial distribution, or it is possible to detect a pattern.

Table A8 – Normality check for provincial random effects in model (5) of Table 3 for R&D

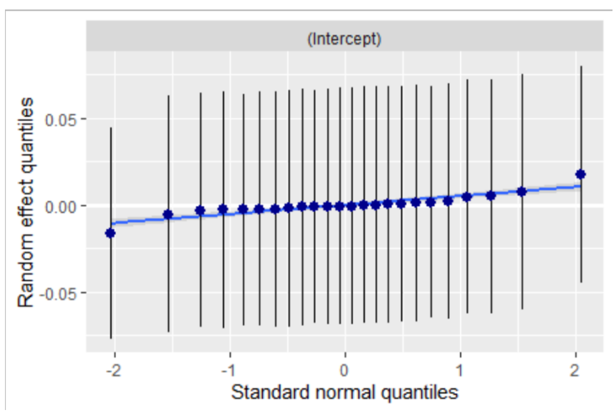
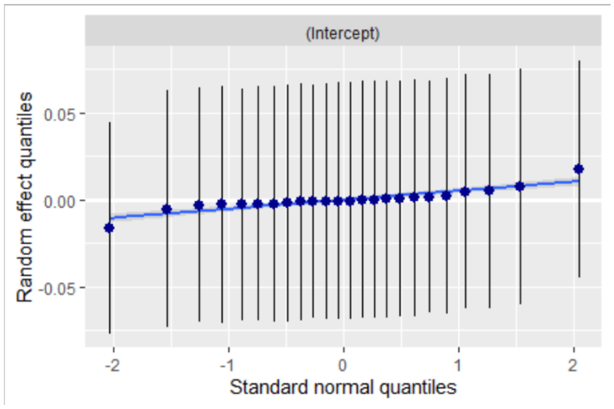


Table A9 – Normality check for provincial random effects in model (6) of Table 3 for Other innovation



We analyze the possible spatial autocorrelation of the estimated random effects using a row-standardized binary weight matrix to the four nearest neighbors. Table A10 shows the resulting Moran´s tests under the hypothesis of normality, comparing the null hypothesis of lack of spatial autocorrelation to the one-side alternative hypothesis of positive autocorrelation. The high p values show that the null hypothesis of lack of spatial autocorrelation is accepted.

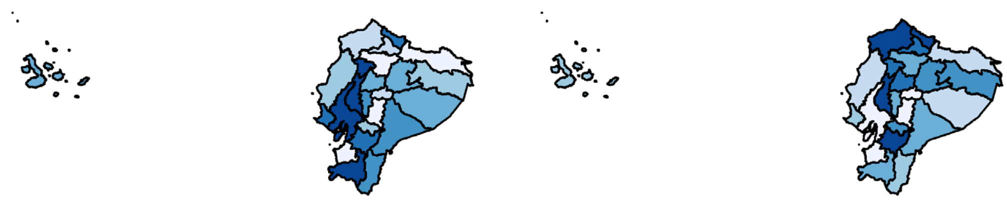
Table A10 – Moran's I of the level-2 residuals of equations (5) and (6) of Table 3

	Moran's I statistic	p value
R&D	0.0538	0.1978
Other Innovation	-0.0907	0.6599

Figure A1 reinforces the previous results showing a map of the level-2 residuals. The value of each provincial random effect estimated in each equation is colored and divided into seven quantiles, which helps to visualize their global spatial pattern. Darker colors are associated with higher values. Though there is some concentration of high estimated provincial effects in a few neighboring regions for the equation of R&D, equation (5) of Table 3, it is not possible to identify a core-periphery pattern in any of the two maps.

Figure A1 – Choropleth maps of the level-2 residuals of equations (5) and (6) of Table 3

R&D **Other innovation activities**



References

Barasa, L., Knoben, J., Vermeulen, P., Kimuyu, P. & Kinyanjui, B. (2017). Institutions, resources and innovation in East Africa: A firm level approach. *Research Policy* 46, 280–91. doi:10.1016/j.respol.2016.11.008

Bellmann, L., K. Evers and R. Hujer (2018). Regional and firm-specific effects on innovations using multi-level methods. *The Annals of Regional Science* 61, 319–49. doi:10.1007/s00168-018-0869-2

Duncan, C., K. Jones and G. Moon (1998). Context, composition and heterogeneity: Using multilevel models in health research. *Social Science & Medicine* 46, 97–117.

James, G., Witten, D. Hastie, T. & Tibshirani, R. (2017). *An Introduction to Statistical Learning: with Applications in R*. Edición: 1st ed. 2013, Corr. 7th printing 2017. New York: Springer.

López-Bazo, E., and E. Motellón (2018). Innovation, heterogeneous firms and the region: evidence from Spain. *Regional Studies,* 52, 673–87. doi:10.1080/00343404.2017.1331296

Neira, I., F. Bruna, M. Portela and A. García-Aracil (2018). Individual Well-Being, Geographical Heterogeneity and Social Capital, *Journal of Happiness Studies,* 19, 1067–90. doi:10.1007/s10902-016-9840-z

O'Sullivan, M. (2005). Finance and innovation. In J. Fagerbeg, D. C. Mowery, & R. R.. Nelson (Eds.). *The Oxford Handbook of Innovation*. Oxford, UK: Oxford University Press.

Snijders, T.A.B., and R. Bosker (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Edición: Second Edition. Los Angeles, Calif.: SAGE Publications Ltd

Sommet, N., and D. Morselli (2017). Keep Calm and Learn Multilevel Logistic Modeling: A Simplified Three-Step Procedure Using Stata, R, Mplus, and SPSS, *International Review of Social Psychology* 30, 203–18. doi:10.5334/irsp.90

Srholec, M. (2010). A Multilevel Approach to Geography of Innovation, *Regional Studies* 44, 1207–20.

——— (2011). A multilevel analysis of innovation in developing countries, *Industrial and Corporate Change* 20, 1539–69. doi:10.1093/icc/dtr024